Pinocchio's Hats

27 June 2022

Jim Stevenson

Wikipedia

This problem in logic from Presh Talwalkar ([1]) recalled an article I wrote a while ago but did not publish. So I thought I would post it as part of the solution.

> Assume that both of the following sentences are true:
>
> - Pinocchio always lies;
> - Pinocchio says, "All my hats are green."
>
> We can conclude from these two sentences that:
> (A) Pinocchio has at least one hat.
> (B) Pinocchio has only one green hat.
> (C) Pinocchio has no hats.
> (D) Pinocchio has at least one green hat.
> (E) Pinocchio has no green hats.

Actually, the question is which, none or more, of statements (A) – (E) follow logically from the two sentences.

My Solution

Symbolic logic interlude. In previous posts on logic problems I translated the language into symbolic logic statements or propositions. These were examples of the "Propositional Calculus".[1] The current problem involves statements about a collection of elements—Pinocchio's hats. Statements where we add quantifiers or variables are examples of the "Predicate Calculus".[2]

For example, as a first pass, the second sentence can be written, where h represents a hat, "For every h, h is green." We have symbols for quantifiers as shown in Table 1 (which includes the other logical symbols for reference). So the second sentence becomes $\forall h$, P(h), where P(h) = "h is green". Now the negation of this statement is "there exists an h, such that h is not green," that is,

$$\sim(\forall h,\ \mathrm{P}(h)) \equiv \exists h \ni \sim\mathrm{P}(h). \qquad (1)$$

(We often drop the $\ni$ after $\exists$, since it is implied.)

Actually, our Pinocchio example is more restricted. We are not saying all hats in the entire universe are green, but just the ones belonging to Pinocchio. So let H be the set of hats belonging to Pinocchio, then sentence two is $\forall h \in \mathrm{H}$, P(h). But what this is really saying is

$$\forall h,\ \mathrm{Q}(h) \Rightarrow \mathrm{P}(h) \ \text{ where } \mathrm{Q}(h) = \text{“}h \in \mathrm{H}\text{”},$$

Table 1

Symbol	Meaning
$\forall$	"for every" or "for all"
$\in$	"in" or "is an element of"
$\exists$	"there exists"
$\ni$	"such that"
$\sim$	"not" or "is false"
$\Rightarrow$	"if …, then …"
$\equiv$, $\Leftrightarrow$	"logically equivalent" or "if and only if"
$\wedge$	"AND"
$\vee$	"OR" (inclusive "or")

[1] https://en.wikipedia.org/wiki/Propositional_calculus
[2] https://en.wikipedia.org/wiki/First-order_logic

that is, "for all hats h, if h belongs to Pinocchio, then h is green.

Now the second sentence is false. So if we let $R(h) = (Q(h) \Rightarrow P(h))$, then from the equivalence (1),

$$\sim(\forall h, R(h)) \equiv \exists h \sim R(h) \equiv \exists h \sim(Q(h) \Rightarrow P(h))$$

Recall that the implication $A \Rightarrow B$, means "if A is true, then B is true". What absolutely must never occur is for A to be true and B false, that is, $(A \Rightarrow B) \equiv \sim(A \wedge \sim B)$. Therefore, the second sentence being false now means

$\exists h\ (Q(h) \wedge \sim P(h))$ = "there exists a hat h, such that h belongs to Pinocchio and h is not green" (2)

Back to the problem. So now consider the possible conclusions (A) – (E), given the formulation (2) of the negative of sentence two.

(A) Pinocchio has at least one hat.	That is certainly true from (2), but it is not green.
(B) Pinocchio has only one green hat.	That may or may not be true, so it does *not* follow logically from (2), since only true statements follow logically.
(C) Pinocchio has no hats.	This is definitely *not* true. (2) says Pinocchio has at least one hat, that is, "there exists an h, such that $h \in$ H", so H cannot be the empty set.
(D) Pinocchio has at least one green hat.	That may or may not be true, so it does *not* follow logically from (2)
(E) Pinocchio has no green hats.	That may or may not be true, so it does not follow logically from (2)

So (A) is the only conclusion that derives logically directly from the two sentences about Pinocchio.

Comment. Conclusion (C) hides a remarkable logical fact. Suppose Pinocchio told the truth, that is, "all his hats are green" is true. Would that still be true if he had no hats? The answer is yes, and it is called a "vacuous truth". That is what my old article is about, which I have included as an Appendix (p.4). Talwalkar also discusses this at some length in his solution.

Talwalkar Solution

Thanks to Guilherme who suggested and translated the problem into English! This problem was asked on the 17th Brazilian Mathematical Olympiad of Public Schools. It went viral on social media because people discussed what the correct answer should be.

We know Pinocchio always lies, so he only speaks mathematically false statements. The statement "All my hats are green" is false if some hat is not green. The statement is definitely misleading—and perhaps a real-world "lie"—if Pinocchio has no hats. These options correspond to answer choices (A) and (C). We will analyze answer choice (C) in more detail below.

(A) Pinocchio has at least one hat.

(C) Pinocchio has no hats.

But first let's examine the other choices one by one.

(B) Pinocchio has only one green hat.	Is this always true? No. Imagine Pinocchio has 2 green hats and 1 blue hat. He could speak a false statement that he has all green hats. Thus choice (B) is wrong.

(D) Pinocchio has at least one green hat.	Imagine Pinocchio has 2 blue hats. He could lie to say he has all green hats. Thus answer (D) is not correct.
(E) Pinocchio has no green hats.	If Pinocchio has 1 green hat and 1 blue hat he would be lying to say he has all green hats. Thus (E) is also wrong.

Most people agree to eliminate (B), (D), and (E). But aren't both (A) and (C) correct answers?

No! From a logical perspective, answer choice (C) is wrong.

Vacuously true statements

A statement is vacuously true if the premise is false or cannot be achieved. For example, consider a room with no mobile phones. If someone says, "All mobile phones in the room are turned OFF," then that would technically be a true statement. After all, there are no mobile phones in the room, so one could say all of the ones in the room are turned off. But for the same reason, one could also speak a vacuously true statement, "All mobile phones in the room are turned ON."

Both statements are technically true, but they are meaningless because there are actually no mobile phones in the room.

Let's apply the principle to the Pinocchio problem.

(C) Pinocchio has no hats.	If Pinocchio has no hats, then any statement he makes about the hats is vacuously true. Thus the statement, "All my hats are green" is technically a true statement. But we know Pinocchio always lies, so this is an impossibility.

Therefore, Pinocchio cannot have no hats; he must have at least one hat. We can eliminate answer choice (C) and the only correct answer is (A) Pinocchio has at least one hat.

References

Pinocchio problem articles

https://ultimosegundo-ig-com-br.translate.goog/brasil/2022-06-08/questao-obmep-pinoquio-viraliza-personagem-mentiu.html?_x_tr_sl=auto&_x_tr_tl=en&_x_tr_hl=en-US&_x_tr_pto=wapp

https://sivtelegram.media/students-make-memes-involving-pinocchio-and-leak-reactions/

https://www.reddit.com/r/askphilosophy/comments/v8brz4/a_question_about_pinocchio_what_is_your_opinion/

https://brainly.com.br/tarefa/52854041

Pinocchio illustration

https://en.wikipedia.org/wiki/File:Pinocchio.jpg

Bram28 explanation vacuously true

https://math.stackexchange.com/questions/2723860/still-struggling-to-understand-vacuous-truths

Wikipedia vacuous truth

https://en.wikipedia.org/wiki/Vacuous_truth

References

[1] Talwalkar, Presh, "Brazilian Olympiad Pinocchio's Green Hats Viral Question", *Mind Your Decisions*, 22 June 2022. (https://mindyourdecisions.com/blog/2022/06/22/brazilian-olympiad-pinocchios-green-hats-viral-question, retrieved 6/27/2022)

Appendix: Vacuous Truth

11 July 2016

Jim Stevenson

I found the following excerpt from a post by Evelyn Lamb to her *Scientific American* blog *Roots of Unity* that mentioned the notion of "vacuous truth" a bit off.

(http://blogs.scientificamerican.com/roots-of-unity/a-few-of-my-favorite-spaces-the-empty-set/, retrieved 7/11/16)

A Few of My Favorite Spaces: The Empty Set

A guided tour of nothing

Evelyn Lamb, June 30, 2016

"Math is about nothing," Andrew Hacker says in a recent New Yorker article[3] by Rebecca Mead. If he had not already,[4] with this statement he has certainly jumped the shark.[5] He follows it up even more perplexingly: "Math describes much of the world but is all about itself," he says. So math is about nothing, except what it is about, which is math and some things in the world? I echo math teacher Patrick Honner, who asks "Why are we listening to Andrew Hacker?"[6]

Hacker's assertion is absurd, but today let's humor him. Meet the empty set.

OK, the empty set is not exactly the most photogenic of sets. It's a set that has nothing inside, and it's kind of hard to get a good picture of nothing. The two most popular depictions of the set are empty brackets { } and something that looks like the Scandinavian vowel ø. (The Wikipedia page for the letter[7] warns us not to confuse the letter with the mathematical symbol, but it does not specify what dire consequences we will face if we stumble.) I will use the symbol ∅ because I think it looks nicer, and when you're writing about nothing, it might as well be well-dressed nothing.

In a recent post on his blog Mathematics Without Apologies, Michael Harris says that ∅ is hardly

[3] http://www.newyorker.com/magazine/2016/06/27/andrew-hacker-debates-the-value-of-math

[4] http://www.slate.com/articles/health_and_science/education/2016/03/andrew_hacker_s_the_math_myth_is_a_great_example_of_mathematics_illiteracy.html

[5] https://en.wikipedia.org/wiki/Jumping_the_shark

[6] http://mrhonner.com/archives/16257

[7] https://en.wikipedia.org/wiki/%C3%98#Encoding

anyone's favorite set,[8] but I think there are a few reasons to give it some love.

Limitless

The first reason to love the empty set is that, far from limiting your possibilities, the empty set opens them up through the magic of **vacuous truth**.[9] Anything is true if you start reasoning from a false premise. We're secretly using this idea rhetorically when we say "I'll enter a digits of pi reciting contest [10] when pigs fly" or "If that's a sensible rail ticket pricing scheme,[11] then I'm the Queen of England."

The empty set is the prototypical generator of vacuous truths. Anything you want to prove about things in the empty set, you can. Logically, this is equivalent to saying "if x is in the empty set, then x has [whatever delightful property you're thinking about]." Do you want a unicorn? Great! Everything in the empty set is a unicorn. Do you hate unicorns? You're in luck—everything in the empty set is a talisman against unicorns. The empty set itself is not a talisman against unicorns, but every last thing in it is.

The empty set is ubiquitous in mathematics, and I mean that literally. It is a subset of every other set. Here, we have to be careful about what we mean by a subset. A set X is a subset of a set Y if every element of X is an element of Y. And when X is the empty set, this is true no matter what Y is! Due to the power of vacuous truth, every element of the empty set is an even number, and every element of the empty set is an odd number. So the empty set is a subset of both the even numbers and the odd numbers. …

Lamb's statement above "Anything is true if you start reasoning from a false premise." seems to focus on one aspect of logic (somewhat misleadingly I fear, since the conclusion could just as well be false) and the statement "The empty set is the prototypical generator of vacuous truths." seems to focus on another, and they don't appear to be the same.

But the fact that Lamb includes both under the notion of vacuous truth is supported by the *Wikipedia* article on the subject (https://en.wikipedia.org/wiki/Vacuous_truth), namely,

> In mathematics and logic, a vacuous truth is a statement that asserts that all members of the empty set have a certain property. … More formally, a relatively well-defined usage refers to a conditional statement with a false antecedent.

Unfortunately, neither Lamb's post nor the *Wikipedia* article provide sufficient details either about the meaning of each interpretation or about why they are essentially the same. I can try to explain each type of "vacuous truth", but I can't really see how they might be considered to employ the same idea. Moreover, it makes more sense to give the name "vacuous truth" to the form using the empty set than the one involving a false premise.

1. Propositional Calculus: P $\Rightarrow$ Q and Truth Tables

This is the "false premise" example of "vacuous truth", namely, the part of the implication P $\Rightarrow$ Q (if P is true, then Q is true) where the *implication* is still true even if P is false and Q is true. This truth comes from the truth table for the implication, namely,

---

8 https://mathematicswithoutapologies.wordpress.com/2016/06/20/is-it-common-knowledge-that-anyone-is-fit-to-be-us-president/

9 https://en.wikipedia.org/wiki/Vacuous_truth

10 http://blogs.scientificamerican.com/roots-of-unity/don-8217-t-recite-digits-to-celebrate-pi-recite-its-continued-fraction-instead/

11 http://blogs.scientificamerican.com/roots-of-unity/british-rail-s-shocking-defiance-of-standard-metrics/

P	⇒	Q
(1)	**(2)**	(1)
T	**T**	T
F	**T**	T
T	**F**	F
F	**T**	F

The reason this is the truth table for the implication is because P ⇒ Q ≡ ~(P∧~Q) ≡ ~P ∨ Q, where ∧ means "and" implying the statement P ∧ Q is true if and only if *both* P and Q are true and where ∨ means "(inclusive) or" or "and/or" implying the statement P ∨ Q is true if and only if *either* P is true *or* Q is true *or both* P and Q are true. The truth tables for ∧ and ∨ are as follows.

P	∧	Q
(1)	**(2)**	(1)
T	**T**	T
F	**F**	T
T	**F**	F
F	**F**	F

P	∨	Q
(1)	**(2)**	(1)
T	**T**	T
F	**T**	T
T	**T**	F
F	**F**	F

Furthermore, we use the logical form of the **De Morgan Laws**: **~(P ∧ Q) ≡ ~P ∨ ~Q** and **~(P ∨ Q) ≡ ~P ∧ ~Q**. And finally **~(~P) ≡ P**.

So what cannot happen if P implies Q is for P to be true and Q false, so either P is false or Q is true or both. That is what the implication means, and that is what the equivalences in P ⇒ Q ≡ ~(P∧~Q) ≡ ~P ∨ Q mean. In fact, the truth tables for these equivalent statements, via the tables for ∧ or ∨, are

~ (	P	∧	(~	Q))
(4)	(1)	(3)	(2)	(1)
T	T	F	F	T
T	F	F	F	T
F	T	T	T	F
T	F	F	T	F

(~	P)	∨	Q
(2)	(1)	**(3)**	(1)
F	T	**T**	T
T	F	**T**	T
F	T	**F**	F
T	F	**T**	F

These tables agree with the truth table for the implication ⇒, that is, the same T/F inputs produce the same corresponding T/F outputs, so the statements must be logically equivalent (see the box below for an explanation of the evaluation steps).

Truth Table Evaluation Steps

The steps in evaluating the truth tables labeled by the numbers in parentheses at the top of the columns are a condensed form for the following procedure, as shown for evaluating ~(P∧~Q):

Step (1) assign all truth values (T/F) to **P** **Q**

Step (2) derive truth value for ~ **~Q**

Step (3) derive truth value for ∧ **P ∧ ~Q**

Step (4) derive truth value for ~ **~(P ∧ ~Q)**

In other words, the implication or "if, then" means if P is true, Q must then be true, and if P is not true, then Q can be anything, either true or false. That is what the implication *means*. So the *implication* is "true" or "valid" even when the statement (premise) P is not. Apparently this has been dubbed a "vacuous truth". But it appears to be more subtle to me and is the source of great confusion whenever logic and the implication are discussed.

2. Predicate Calculus: $\forall x \in \varnothing, P(x) \equiv \sim(\exists\, x \in \varnothing \ni \sim P(x))$

Now turning to the effect of the empty set, we have "For every x in the empty set $\varnothing$, the statement $P(x)$ is true" is logically equivalent to "It is false that there exists an x in the empty set $\varnothing$, such that the statement $P(x)$ is false." In symbolic logic this is written $\forall x \in \varnothing,\ P(x) \equiv \sim(\exists\, x \in \varnothing \ni \sim P(x))$, where $\forall$ means "for every" or "for all", $\in$ means "in" or "is an element of", $\equiv$ means "is logically equivalent to" or "if and only if", ~ means "not" or "it is false that", $\exists$ means "there exists", and $\ni$ means "such that" (which is often omitted and implied implicitly after $\exists$). The right hand side is definitely true, since "there exists an x in the empty set $\varnothing$, such that the statement $P(x)$ is false" is definitely false because there cannot exist any x in the empty set $\varnothing$.

Perhaps I am being too casual here. First I need to make explicit a type of De Morgan Law for predicates, namely, $\boldsymbol{\sim(\forall x,\ P(x)) \equiv \exists\, x \sim(P(x))}$ ($\ni$ has been omitted for simplicity). This means the statement "it is false that for every x, $P(x)$ is true" is equivalent to "there exists an x, such that $P(x)$ is false."

Furthermore, the meaning of $\forall x \in \varnothing,\ P(x)$ is really $\forall x,\ x \in \varnothing \Rightarrow P(x)$, which is of the form $\forall\, x,\ Q(x) \Rightarrow P(x)$. Therefore its negation is $\sim(\forall x,\ Q(x) \Rightarrow P(x)) \equiv \exists\, x\ \sim(Q(x) \Rightarrow P(x)) \equiv \exists\, x\ Q(x) \wedge \sim P(x)$. But $\exists\, x\ Q(x)$ = "$x \in \varnothing$" is true, is a false statement, so the whole statement must be false, namely, the statement "$\sim(\forall x \in \varnothing,\ P(x))$ is true" must be false, so that "$\forall x \in \varnothing,\ P(x)$" must be true!

In any case, I do not see these two examples as being the same. It is true that an implication is involved with a false condition and true conclusion, as in the previous case, but the key here is $\exists\, x$, "there *exists* an x." The predicate is essential. So I am surprised they are both called vacuous truths (the moniker seems to fit the second case involving the empty set, but not the first only involving a false premise).