

Classic Paradoxes Unraveled

Tony Polito, Ph.D.

There are a number of famous classic paradoxes ... that aren't really paradoxes at all. And have quite rational explanations. I present some of those supposed paradoxes and their resolutions below.

A true paradox is where an argument, developed from some set of underlying assumptions and/or premises, results in a clearly false conclusion ... or develops two lines of formal logic resulting in directly opposing conclusions. Such a true paradox is considered proof that the underlying assumptions and/or premises upon which the argument is based must be false. In formal logic this is described as *reductio ad absurdum* (or *RAA*), a Latin phrase implying "the argument has reduced to absurdity." So, in essence, there is nothing really wrong with a true paradox; rather, *RAA* brings a true paradox to a very logical resolution.

Most classical paradoxes are not actually paradoxes in this true sense, but only appear to be. Below I unravel the true nature of these supposed paradoxes and reveal their resolutions.

The Raven Paradox

(authored October 13, 2019)

Philosopher Carl Hempel presented this supposed paradox in the 1940s in an attempt to demonstrate that there can be inconsistency between formal logic and deduction by intuition/common sense.

Somebody presents a hypothesis: All ravens are black.

We reword that slightly, so as to make it a more formal conditional statement: If it is a raven, then it is black.

We can even write the statement using the symbols of formal logic should we wish: $R \rightarrow B$, where *R* stands for raven, *B* stands for black and the arrow represents the "if-then" conditional relationship.

This statement, like many hypotheses, can never be proven to be true. If we examined every raven that ever lived, or is living, and they all were black, all those occurrences would certainly serve to increase the likelihood that the hypothesis is true. But we will never know if the future might bring us a raven that is not black. We can only *falsify* the hypothesis: show me just one raven that is not black, and the hypothesis is proven to be wrong ... and logically worthless.

Said another way:

My neighbor stops by. He has a raven and it is black. This occurrence *does not* falsify, though it serves to increase the likelihood that the hypothesis is true. Then another neighbor stops by. He has a raven and it is red. This occurrence *does* falsify; the hypothesis is incorrect. We can express it

symbolically as $(R \rightarrow B) \wedge R \wedge \sim B$, (the \wedge means "and" and \sim means "not"). That is a statement of symbolic logic that will always be valued as false.

One rule of formal logic that is always correct is that the contrapositive restatement of a conditional statement is equivalent to the underlying conditional statement. A contrapositive restatement 'flips' the two entities and changes them to "not." Regarding the raven hypothesis, the contrapositive statement is "If it is not black, then it is not a raven." Of course! That's just restating the hypothesis in a slightly different way: it couldn't BE a raven unless it WAS black! Symbolically that contrapositive statement can be written as $\sim B \rightarrow \sim R$.

Now we arrive at the supposed paradox. Suppose I step out to my garage where my car is parked. My car is red. According to (the contrapositive of) the raven hypothesis, since it is not black, it is not a raven. True enough.

So, based on the above discussion, this occurrence serves to increase the likelihood that the raven hypothesis is true. But wait a minute! Intuition and common sense tells us that the color of my car doesn't have anything to do at all with ravens or their color!

Therein, supposedly, lies the paradox. *In terms of formal logic*, by looking at my car in my garage, I now know more about ravens and their color. *In terms of intuition & common sense*, looking at my car color couldn't possibly teach anything new at all about ravens and their color.

So there is the inconsistency of which Hempel speaks. But that does not make it a true paradox. Again, a true paradox is where an argument developed from same set of underlying assumptions and/or premises results in a clearly false conclusion ... or develops two lines of formal logic resulting in directly opposing conclusions. What we actually have here is an example of correct logic correctly trumping incorrect intuition & common sense. The way the raven situation is described/framed, it leads our intuition & common sense astray ... when the logic is actually correct.

Looking at a different example clarifies.

My Aunt Maria is full Italian and a devout Catholic. Devout Catholics still don't eat meat on Friday. So I believe she will serve fish for dinner every Friday. But I believe that every other night of the week, she will serve up an old-school, hearty Italian meal. I stop by for dinner on Friday. I see fish on the dinner table.

The hypothesis here is "If it is Friday, dinner is fish." This occurrence does not falsify and so serves to increase the likelihood that the fish hypothesis is true.

However, we could *claim* the same "paradox" exists as well. *In terms of logic*, we now know more about when fish is served. But couldn't we also say, *in terms of intuition & common sense*, we couldn't have possibly learned anything about what's going to be on Aunt Maria's dinner table when I get there ... just by looking at a calendar! But we don't say that. Because *this* relationship, unlike a red car and ravens, seems to make much more intuitive sense, there is a rational 'connection.' So we are not troubled by the logic of it.

In the fish example, we described/framed the situation with some reasonable rationale as to why the two entities might well be closely related. So our intuition & common sense 'accepts' the notion that we've actually learned something about the fish hypothesis. Our logic and our intuition & common sense are in agreement.

The raven example is described/framed in such a way to ***deceive*** our intuition & common sense to ***not*** agree with the logic.

Suppose I re-framed the raven example slightly and said: "The bird-keeper at the local bird-zoo has developed a hypothesis that all ravens are black. He spends the day showing me all the birds in the zoo. He shows me a bird. It is red and it is a robin."

Said this way, our intuition & common sense sees some reasonable rationale as to why the two entities might well be closely related. Instead of talking about "all things in the world" and "ravens" now we are talking about "all types of birds" and "ravens." Our intuition & common sense sees a reasonable relationship between "all types of birds" and "ravens." So when shown a red robin, we see that occurrence as actually having taught us something about the validity of the raven hypothesis. Our logic and our intuition & common sense are in agreement.

That being said, it doesn't actually make any difference whether it was a red car or a red robin. In each occurrence, we were NOT shown "a raven that is not black." So the occurrence did not falsify the hypothesis and so it served to strengthen the raven hypothesis.

Now one might argue looking at car colors doesn't add as much strength to the raven hypothesis as does looking at bird colors. To an extent, that is true. But again, a million black ravens observed does not preclude an albino raven tomorrow. In that sense, a hypothesis is strong only because it has not (yet) been falsified, more so than from all the occurrences that did NOT falsify it. And once a hypothesis is falsified, it is worthless. So a red car occurrence contributes as much strength to the hypothesis occurrence as does a red robin. Because, in each case, they did not falsify.

In fact, were an observer to collect a million red robin occurrences, he/she might come to feel that yet another red robin occurrence would add no strength whatsoever to the hypothesis, just as he/she might feel about the occurrence of the red car in my garage. But, again, the truth of the matter is that they the both contribute the same strength to the hypothesis. Because they did not falsify.

The original raven example is stated/framed in a way so as to mislead our intuition & common sense into not agreeing with the logic ... but the logic is quite correct.

The Unexpected Hanging

(authored October 28, 2019)

My first exposure to this supposed paradox was back in the 1970s from a book by that title authored by Martin Gardner. Gardner authored many books on mathematical, scientific and logical curiosities that were drawn from a column he published in *Scientific American* for over twenty years.

One Monday morning, a judge hands down his sentence to a convicted criminal. "You will hang by the neck within the next seven days. And, to enhance your suffering, I will assure you do not know the day of your hanging until it arrives."

When the criminal returns to his cell, he is actually pleased! The criminal thinks to himself:

"The hanging cannot be on Sunday. For if Saturday comes and goes, I will know the hanging will be on Sunday. The judge, a man of his word, therefore cannot hang me on Sunday."

"Further, the hanging cannot be on Saturday. For if Friday comes and goes, I will know the hanging will either be on Saturday or Sunday. But the hanging cannot be on Sunday. For if Saturday comes and goes, I will know the hanging will be on Sunday. Therefore if Friday comes and goes, I will know the hanging will be (not on Sunday but) on Saturday. The judge, a man of his word, therefore cannot hang me on Saturday (either)."

The criminal extends his logic along these lines until he has logically concluded that he cannot even be hung on Monday. That he cannot even be hung at all! The hangman will never come!

Then, on Wednesday, the door to the criminal's jail cell swings open. It is the hangman, come to carry out the sentence. The criminal, having already convinced himself he cannot be hung at all, is completely taken by surprise. True to the judge's word, the criminal, confused by his own logic, did not know the day of his hanging until it arrived.

This too is not truly a paradox. Again, a true paradox is where an argument developed from same set of underlying assumptions and/or premises results in a clearly false conclusion ... or develops two lines of formal logic resulting in directly opposing conclusions.

Suppose the criminal's thinking had been somewhat different:

"That judge will never hang me. He knows I am a member of a violent gang and that if he hangs me my gang will kill him and all of his family."

Or perhaps:

"That judge will never hang me. He is devoutly religious and so he believes that if he hangs me his soul will perish in Hell for all eternity."

Or perhaps:

"That judge will hang me on Friday. He has been hanging criminals for twenty years and he always hangs them on Friday. He said I wouldn't know, but he's a liar. Everybody knows he hangs 'em on Fridays."

Now all those lines of thinking are flawed. Whether the criminal thinks they are logically sound or not. Some judges are fearless when it comes to incarcerating ... or condemning to death ... violent

gang members. Some judges are devoutly religious ... and yet they still sentence criminals to death. And just because the judge has always hung criminals on Fridays in the past does not absolutely assure he will continue to do so in the future.

So in these cases we easily see that the criminal's lines of thinking do not have any actual influence or impact upon what will actually happen.

And therein is the shortcoming.

The criminal's thinking ... about not being hung on Sunday and such ... does not have any assured impact on what will actually happen.

The (original) line of logic in the criminal's mind that is consistent with an impossible hanging does NOT fall from the same set of conditions as the arrival of the unexpected hanging that falls from a set of conditions in reality. Said another slightly different way, the conclusion of the impossible hanging is based upon the conditions that the criminal knows in his mind, while the *actual* arrival of the unexpected hanging is based upon on the conditions that the criminal ends up *really* knowing "in the real world."

In fact, in general, there is no assured and equivalent connection between what anyone knows about reality in their mind and their own actual reality.

In the real world, the judge can pick any day he wishes for the hanging and, if no one tells the criminal, well the criminal *really* doesn't know. In the real world, if Friday comes and goes, the criminal *really* has no way of knowing whether he will hang on Saturday or Sunday, if no one has told him one way or the other. The criminal may logically deduce in his mind whatever he thinks he knows about the matter, but it will have no influence whatsoever on what day the judge *actually* chooses or what the criminal *really* ends up knowing. If the criminal is still un-hung on Sunday, well *then* he will know what day he will hang, but then again, now that day has arrived.

In the discipline of philosophy this issue is discussed as the topic of "externalism." That is, to what extent does what one knows in their mind equivocate to what is truly real. Are they the exact same thing? Well the answer is that they *can* be, but there's no guarantee that they *will* be. And there's no way for a mind to be sure they *are*.

They *can* be. Mental formal logic can certainly lead to correct conclusions regarding reality. Doing geometric logic in one's mind can correctly expect the measurement of actual angles and sides of triangles. Mental logic can often correctly predict the outcome of experiments in chemistry or physics.

But there's no guarantee that they *will* be. Many scientists have mentally deduced what they believed to be logical conclusions about chemistry or physics, only to be eventually proven wrong.

And there's no way for the mind to be sure they *are*. The most extreme example that a reader would be readily familiar with would be the plight of Neo in the film *The Matrix*. At the beginning of the film, in his own mind, Neo knows without any doubt (or reason to doubt) that he is a computer programmer by day and a dealer in illegal software by night, a crime

that causes him to be pursued by government agents. Of course none of Neo's 'knowledge' is even remotely correct, as the audience later discovers when it is revealed that Neo's perceptions are being fed into his brain via technology ... while he *really* exists as a nameless unit in a giant human battery farm. In the discipline of philosophy, this exact concept has long been known, condensed into a scenario known as "the brain in the vat" (a concept upon which this film clearly draws). How could a brain suspended in a vat of fluid that keeps it alive, while it is wired to a computer feeding it perceptions of a robust life in a human body, how could that brain ever know it wasn't anything more than a brain in a vat? How could Neo possibly know (without any external influence such as Morpheus) that he is not really anything more than a human battery? The answer is that they cannot. This theme of "how can our mind know what is real" was a favorite of sci-fi writer Philip K. Dick. It appears in films derived from his written works such as *Total Recall* and *Minority Report*.

The criminal's logic, derived from the set of conditions within his own mind, about how the judge must go about selecting the day of the hanging, may appear sound. But there's no assurance that his mental assumptions or logic will play out in the same manner outside his mind, in reality.

And, in fact, they do not. The *real* process of selection and knowing has nothing to do with the process of selection that the criminal believes to 'know' from his logical deduction.

The Paradox of the Court

(authored October 31, 2019)

This paradox originates from ancient Greek philosophy and literature.

A teacher of law took on a student of law. A contract was written making payment for the instructional services due only after the student has won his first case in court.

The student of law never entered the profession of law whatsoever ... and so he did not pay the teacher. The teacher sued in court for the amount owed.

The teacher argued that (1) his winning of the case is a decision that the student must pay the money. And that (2) his losing of the case means the student is obligated to pay the money now, given the terms of the contract that the student must pay after winning a case.

The student argued that (3) his winning of the case is a decision that he does not have to pay the money. And that (4) his losing of the case means he is *not* obligated to pay the money yet, given the terms of the contract that the student need not pay until winning a case.

So who is right, student or teacher?

While this appears to be a paradox—in that both teacher and student have each used the same set of conditions to logically reach opposite conclusions—truly it is not. For what we have here is, upon closer inspection, is simply poorly-formed logic.

The flaws within the crafting of the logic are the confounding of the terms of the contract and the court decision ... as well as the confounding of the timing/sequence of events.

This can be seen most easily in Statement 4.

If the student loses the case, then the decision is that he must pay. Period. The contract does not relieve him of that decision. Indeed the purpose of the case in court is to determine whether or not the terms of the contract will apply. The logic expressed has incorrectly 'mingled' the dependencies of the decision and the terms of the contract. (The same logical flaw exists, in reverse, within Statement 2.)

Two parties that go to court concerning their contract are usually there because they interpret their contract in two different ways. They present their arguments, then the judge reaches a decision regarding the correct interpretation of the contract. And it is *that* interpretation that is enforced. Neither student nor teacher can use his own interpretation of the contract requirement to *require* a decision in his favor.

And Statement 4 has also disordered the timing/sequence of events. The student is arguing he should win his case (now) based upon conditions that (he believes) will occur (later) if he does not win (now). Again, the same logical flaw exists, in reverse, within Statement 2.

As a practical matter, here is what would likely transpire (unless the matter plunged deeply into historical precedence set in contract law):

The terms of the contract are unambiguous. The student does not owe the money until he wins his first case. And, at the moment the case is being heard, the student has not (yet) won his first case.

Further, the teacher really can't argue the circumstance that he is *absolutely* owed the money, since he contracted to terms where there was significant risk of non-payment; the student, for example, might *never* win a case ... or the student might pass away before he does.

The decision would be in the student's favor. The student has not (yet) won a case and so he need not pay.

The next day, however, the teacher may accost the student for payment (again), according to the terms of the contract. The student is likely to reply that the decision has already been made, that he does not owe the money, and refuse to pay. If the teacher takes the student to court (again), the judge will likely find that the student has (now) won his first case and so must (now) pay the teacher.

Again, the original Statements 2 & 4 are logically flawed and so they are invalid argument. And Statements 1 & 3 say nothing that is not already known before entering the court. So the truth of the matter is that neither the teacher nor the student actually presented any kind of viable argument at all to the court, much less a paradoxical one.

The Paradox of The Heap

(authored November 12, 2019)

This paradox also originates from ancient Greek philosophy and literature.

"There is a heap of sand. Take away a single grain of sand. It is still a heap. Take away another single grain of sand. It is still a heap. Because one grain of sand does not make the difference between a heap and not a heap. Take away a another single grain of sand. It is still a heap. Eventually, there are only two grains of sand. Take away one of them. What is left is still a heap. Take away the last grain of sand. Though all the grains of sand are gone, there is still a heap."

This paradox can also be described in a reverse sequence.

"We add a single grain of sand to another grain of sand. This surely does not make a heap. We add another single grain. Still, not a heap. Because one grain of sand does not make the difference between a heap and not a heap. Add as many grains of sand as you like, one at a time. There will never be a heap."

Perhaps even more frightening to serious philosophers, this line of thought can be used to deny the existence of ... or instantaneously materialize from thin air... any object.

"There is a car parked in front of your residence. It is made of a heap of atoms. We use technology to remove a single atom. It is still your car." As with the heap, eventually there are no atoms, but your car, logically, must still be there. Even though it is not. Or, working in reverse, add all the atoms you like starting with the first atom, but your car will never be there, even if you add all the atoms required.

There does not appear to be any agreed-upon resolution to this paradox amongst philosophers.

A typical response among neophytes to this supposed paradox is that it is caused by *vagueness*, that "heap" lacks adequate precision in definition. But that does not solve the problem. Suppose a "heap" is defined as 100,000 grains of sand. Then 99,999 grains of sand is not a heap. So if you take a single grain of sand away from 100,000 grains of sand, then you now no longer have a heap. Under that definition of heap, you know exactly when you have a heap and when you do not. However, were we to put two piles of sand side-by-side, one pile with 100,000 grains, the other with 99,999 grains, most people would find it very difficult to stand on the notion that one pile is a heap while the other is not. Why? Well because, as said above, one grain of sand does not make the difference between a heap and not a heap.

Another response that attempts to remove the issue of vagueness, is multi-valued or 'fuzzy' logic. As the paradox is described, there are only two truth values available: heap or not-heap. So the problem is approached by employing a type of logic where more than two truth values are available. Suppose we incorporate three truth values: (1) absolutely a true heap, (2) absolutely not a true heap ... and (3) in the middle, sort of a heap, but not really a true heap. That seems to allow a process of transition from heap to not-a-heap, that there is a point where it can be said that the pile is not really a heap anymore, but it's not really not-a-heap either. But that approach just brings us back to the original

dilemma. Why? Well because one grain of sand does not make the difference between 'absolutely a true heap' and an 'in the middle sort of a heap.'

The correct starting point for resolution is to recognize that the formal logic is flawed. "A given" minus 1 does NOT equal "a given." $X - 1 \neq X$. That, from a position of formal logic, is undeniable. So it is not formal logic that is providing the truth value of 'still a heap' after the removal of a single grain of sand. Instead, what is occurring (incorrectly) is that the truth value of 'still a heap' after the removal of the single grain of sand is first being concluded by the actor, then that conclusion is being subtly and intuitively applied to create the illusion of apparent logical veracity of a statement that is actually logically false.

What appears to be a line of logical reasoning leading to a false conclusion, a paradox, is actually the incorrectly blending of truth values derived from formal logic and the truth values being applied by the actor.

Consider this statement: "There is a heap of Democrats in America." Given about 50% of Americans vote Democrat, I expect that Republicans and Democrats would agree this statement is true: Republicans would claim they are working to overcome a heap of significant Democratic force at election-time; Democrats would say they represent a heap of Americans that don't agree with Republicans. Now let's tweak that statement slightly. Instead of stating there are 'many' Democrats in America, let us state "There are TOO many Democrats in America." Now the truth values applied will diverge, dependent upon the particular actor: Republicans will say that statement is true, that Democrats are the problem and we need less of them, while Democrats will say that statement is false, that more Republicans need to change their ways and cross-over to the beliefs of the Democratic Party.

The above example makes clear that the truth values of 'heap/many' and 'too many' are highly dependent upon what is applied by the actor.

But the fact of the matter is that the truth value of *any* logical statement is, to some extent, dependent upon the actor. Formal logic can *never* claim its truth values are *entirely* disjoint of the truth values applied by the actor. It is merely a matter of *degree*.

Consider this simple example where logic is dependent upon the actor. "If a year has passed, the Earth has rotated around the Sun exactly one time." From where does the truth value of the entire statement derive? It is derived from the actor. Time-wise, it takes 365.256 days (a solar year) for the Earth to rotate around the Sun exactly one time. (That's why we have a leap day every four years, to 'catch up' our calendar.) One actor might think that the phrase 'a year has passed' refers to a solar year, and so derive a value of 'true' for the entire statement. Another actor might think that the phrase 'a year has passed' refers to a calendar year, 365.000 days, and derive a value of 'false' for the entire statement. While the logic is sound either way, the truth value of this simple statement is clearly dependent upon the actor.

And so it goes for all truth values in formal logic. In many, many cases of argument by formal logic there is high *consensus* among any and all actors, in that most everyone agrees upon the truth value of the statements, a high-enough degree of agreement that the statements can be said, as a practical matter, to be absolutely true or absolutely false. And so the argument can be adequately evaluated. In some cases, statements can be re-written so as to exclude incorrect injections of truth value; we

could have been more specific about the term 'year' in our solar example. However, in many other cases, the degree of consensus may not be sufficient to provide a truth value that can correctly evaluate the truth of the entire argument.

What has 'gone wrong' in the heap paradox is not a matter of vagueness. Many statements will contain some, if even an insignificant amount, of vagueness. Rather it is a matter of a lack of sufficient consensus. While most actors will agree that one less grain of sand is *not enough* (to change a heap to a not-heap), there is almost no consensus as to what *is enough*, what WILL change a heap to a not-heap. So actors in general, nudged by the logically incorrect statement that 'a heap minus one is still a heap' will (incorrectly) apply what consensus they do possess, apply what they *are* certain of, that one grain is *not enough*, and so apply truth (incorrectly) to the statement 'still a heap.' And that leads actors to conclude (incorrectly) that the entire 'a heap minus one is still a heap' statement is logically consistent. When, plainly, it is not.

The argument presented in the heap paradox is not logically correct; 'a thing' take-away something is NOT still the same thing. That in itself means this is not truly a paradox. The lack of consensus regarding "what a heap is" is causing the incorrect injection by actors that 'still a heap' is true, leading to the incorrect conclusion that the entire 'a heap minus one is still a heap' statement is true, and then to proceed to evaluate the entire argument as if it were logically consistent. When it is not. And the logically inconsistent argument is leading to a conclusion that makes no sense.

The Monty Hall Problem

(authored November 20, 2019)

Though this often gets thrown into the category of paradox, it too isn't really a paradox at all. Rather it is correct logic that results in a correct, though highly un-intuitive, conclusion. And that conclusion is now widely agreed upon. That being why it more commonly called "The Monty Hall Problem."

I discuss it here primarily to try to present an explanation of the conclusion that makes it easier to see why it is correct.

The Monty Hall problem first gained notoriety in 1990 when it was published in a nationally syndicated newspaper column. However Martin Gardner described a similar problem as far back as 1959. It is named for the original host of the "Let's Make A Deal" television show within which the described scenario frequently occurred.

A contestant is given the chance to choose between three closed & numbered doors. Monty informs the contestant that behind each door is a prize he/she will win, that behind two doors is a goat, behind the other door is a brand new car. The contestant picks a door. Then Monty, who knows what is behind all three doors, opens one of the two doors not picked, and there stands a goat. Monty uses the goat-reveal to tempt and/or confuse the contestant, asking if he/she wishes to keep the door picked ... or switch to the other unopened door. Should the contestant keep or switch?

Surprisingly, the contestant *should* switch; doing so doubles his/her chances of winning the car (versus keeping the original door picked). It's absolutely true ... but seriously counterintuitive. I

myself did not believe the solution to be correct when I first saw it, not for quite some time. However, it has been proven mathematically, computer simulations of many thousands of trials have been run and the results confirm the solution. And, get this, pigeons being rewarded with food for correct decisions learn rather rapidly to switch. You can even find webpages where you can play this game over and over, and watch the tallies and percentages accumulate as you use a 'keep' strategy or a 'switch' strategy ... and watch the numbers prove that switching is better, right before your own eyes.

What seems intuitive is that the opportunity to switch is meaningless; there are two doors, the car could be behind either one, so there's a 50-50 chance it is behind one door or the other. So it doesn't matter either way.

Intuitive, but incorrect. Let's break it down.

At the beginning, before the contestant picks any door, the chances are quite clear. The contestant has an one-third chance of picking the correct door. That means there is a two-thirds chance that the car is 'with the rest of the doors.' Now it's too bad that Monty did not offer the option to pick 'the rest of the doors' instead of just one door. Because that would have been a better bet, with twice the chance of winning the car. But he didn't offer.

Monty then opens one of the other two doors, reveals a goat, then makes his offer, keep or switch?

Well, consider now, what has really changed so far?

The answer is 'not much.'

- * The goats and the car are still exactly where they were at the beginning.
- * The contestant's odds haven't changed. There were three doors available and the contestant picked one of them. That's a one-third chance of being correct. Still.
- * The contestant already knew that at least one of the other two doors had a goat behind it, no matter what, so Monty hasn't really provided any new information.

What has changed is that Monty is now offering that better deal, to pick 'the rest of the doors!' If the contestant switches, he/she is swapping out his/her original pick for either (1) a car and an already opened door or (2) a goat and an already opened door.

Again, it doesn't really make any difference that Monty showed one of the door is a goat; the contestant already knows at the outset that, for any two doors, at least one of the doors is going to ultimately reveal a goat.

What does make a difference is that Monty is now offering the chance to swap out the one door for 'the rest of the doors.' Two doors, the chances still are two-thirds that the car is 'with the rest of the doors.' Just like it was at the beginning.

This line of thinking is more intuitive if we change the problem so that number of doors is 100. Let's say you are the contestant and you pick one of the 100 available doors. That's only a 1%

chance of being right. You can be pretty darn sure, 99% sure, that the car is somewhere 'with the rest of the doors.' You can only hope that your pick was a stroke of luck, like with a lottery ticket, that the door you picked hides the car. Then Monty opens 98 of the 99 other doors ... and reveals 98 goats. You 'already knew' the car was (almost certainly) 'with the rest of the doors' so now you are pretty darn sure that car is behind that 99th door Monty didn't open. It's not a 50-50 deal ... because all you (still) have in-hand with the door you picked is that hope that your pick was a stroke of luck, like with a lottery ticket, that the door you picked hides the car.

So when Monty offers, you switch.

And, when you switch, your odds change from 'almost nothing' to 'almost a sure thing.'

The Ship of Theseus

(authored December 15, 2019)

Perhaps the oldest documented supposed paradox, first described by the Greek historian Plutarch in his *Life of Theseus*. Theseus was the mythical heroic King of Athens, slayer of the Minotaur of the Labyrinth. When he returned to Athens to take the title of King from his father King Aegeus, his ship was preserved in the Athenian harbor as memorial for several centuries:

"... they took away the old planks as they decayed, putting in new and stronger timber in their place, in so much that this ship became a standing example among the philosophers, for the logical question of things that grow; one side holding that the ship remained the same, and the other contending that it was not the same."

So the question is, once all the parts of the Ship of Theseus were replaced, could it still rightfully be called The Ship of Theseus?

While this may sound the stuff of abstract philosophical debate, we are confronted with similar, quite pragmatic, quandaries in every-day life:

- A professional sports team, over the years, moves from one city to another, yet another. Over that time, all of its coaches, players and employees have completely changed several times. Yet people still speak of it by the same name: The Raiders—Oakland, Los Angeles, Oakland, Las Vegas; The Rams—Cleveland, Los Angeles, St. Louis, Los Angeles again. Haven't they stopped being The Raiders and The Rams after all that? On the other hand, many teams have moved from one city to another AND immediately changed the team name, even though the team essentially has all of the same coaches, players and employees. And the new name is readily adopted by most everyone. Why is it that it is not the same team anymore?
- The same with musical bands. Over the years, members come, members go. When does it stop being the original band? The 1960s soul musicians known as The Drifters. All the members left to form their own band, but the owner of the legal rights to the name hired new band members to replace them. Which is truly The Drifters, The Drifters ... or the deserters who took the name The Original Drifters? Over time, more band members in both of them come and go, pass away, till none of those from the very beginning remain in either. The ownership of the

Drifters name passes on to others from the estate of the original owner. Now which is truly The Drifters? Both? One? Neither?

- Rivers. If you swim in the Mississippi River this year, then again next year, did you swim in the same river? A river is water and none of the water is the same. Over time, The Mississippi River changes its course, one bend of the river dries up while the river bends the other way. But we still call it the Mississippi River. All this is true of the Colorado River ... and more, for the Colorado River, once flowing over the surface, no canyon, has slowly cut a path through stone until it is at the bottom of a canyon a mile deep. And someday it will be at the bottom of an even deeper canyon. Is it still the same river, though it never has the same water, path or even height?
- A man is born a baby, then he is a child, then an adult, then an elder. Never of the same size, height, weight or appearance. Never of the same intelligence, maturity, memories or mind. Over time, every cell in his body has been replaced, most many times. What thing is it about this man that causes us to continue to call him by the same name? In his youth he may have been a hooligan, even sent to jail for it, then only to soften and repent with age. But we call the hooligan and the repentant by the same name. Even when he is dead and gone, when he no longer actually exists, people will speak of him in the present tense by the same name, as though he still exists. And the same name will appear on his tombstone, though there is nothing under it but indistinguishable ashes and dust.
- Modifications to the puzzle of The Ship of Theseus. All the wood parts are saved and reassembled in a warehouse. Is the ship in the harbor or the ship in the warehouse the true Ship of Theseus? Or are both The Ship of Theseus? What if all the wood parts of the ship in the harbor are replaced by durable metal parts. Could the ship in the harbor still be reasonably called The Ship of Theseus?

The puzzle of The Ship of Theseus is bottomless. It seems to reveal that there is no logical relationship between what we call a thing ... and what a thing actually is.

At first blush, The Ship of Theseus seems similar to the problem posed by The Heap. Take away a single grain of sand, is it still a heap? Take a single original plank of wood, is it still The Ship of Theseus?

But they are not the same. In the former, we speak of whether the thing still belongs in its category, a matter of definition. In the latter, we speak of whether the thing is still the same thing, a matter of identity.

Indeed, the conundrum that surrounds The Ship of Theseus (and the teams and the bands and the man) is that of identity.

The resolution is actually rather simple. Logic does impose constraint upon what constitutes membership of an entity within a set of entities (that comprise a category); the entity must possess the characteristics required to be a member of the set as it is defined. However logic does NOT impose any constraint upon how the human mind chooses to assign identity to an entity. We can continue to call a sports team by its original name as long as we wish, until we deem the original

name no longer sufficiently relevant to represent it going forward. And that relevancy can be deemed differently from one person to the next. The Colorado River is in the same general area, follows the same general path and does what a river generally does, so most everyone chooses to continue to call it The Colorado River in reference to its past, present and future. A man is born with a name, no other man ever physically replaces him, so most everyone chooses to continue to call him by his original name. On the other hand, when a place is renamed, some will rapidly adopt the new name and the significance the new name represents, others will continue forever to speak of it by its original name, the name associated with its history. For some it is Reagan National Airport, for others it will always be 'National Airport.' For some it is Jackson-Hartsfield Airport, for others it will always be 'Hartsfield.' Yet no one chooses to call New York City by New Amsterdam any more, though that is its original name and its history. The assignment of identity is, ultimately, a matter of choice, not a matter of logic.

The reason that The Ship of Theseus puzzle perplexes us when we first read of it ... so much more than bands or teams or airports ... is that both choices seem very, very equally reasonable. And so there is solid, evenly-divided disagreement ... between all those philosophers ... and also between all of us. But there is nothing paradoxical or incorrect about that. For each human mind can choose to assign identity to an entity as it sees fit, unconstrained by any rules or any logic. Over time, perhaps minds will reach a consensus about the identity of a particular thing, or perhaps not. But there is no constraint imposed by logic that requires it. Each person can view the identity of The Ship of Theseus as he/she sees most appropriate.

Tragedy of the Commons & Prisoner's Dilemma

(authored December 20, 2019)

These are not usually referred to as paradoxes, yet they are similar to some (supposed) paradoxes in that they describe logical decisions that result in sub-optimal outcomes.

Hunting has brought whales ever-closer to extinction since the late 19th Century. By the middle of the 20th Century, the ever-improving efficiency of whale-hunting methods and the massive whale-processing ships brought extinction so close that in 1986, all country members of the International Whaling Commission, such as Norway, Iceland, Japan, South Korea, agreed to a total ban on whale hunting. Yet Japan continued to whale hunt, albeit on a smaller-scale, under the thin guise of clauses in the agreement allowing exclusions for scientific purposes and indigenous peoples. Japan continued to push for the ban to be lifted, allowing for annual whale hunting quotas, without success. So in 2019 Japan left the IWC and resumed full-scale, unrestricted whaling. Lacking the cooperation of Japan, other countries are likely to do the same soon enough.

It would be in the best interest of all countries to cooperate about whaling so that the population can reach and maintain a sustainable level, ensuring the never-ending availability of whales. But instead individual countries continue to pursue whaling at a pace that can and likely will cause eventual whale extinction, depriving all of them of the opportunity for further whaling forever. The question is why do they not cooperate when it would be in all of their best interests?

This type of scenario is typically described as a "tragedy of the commons," first described as a problem with the overuse of common grazing lands in the 1800s. Modern mathematical analysis in the form of game theory¹ rationally explains what appears to be, at face-value, the illogical behavior.

Let us reduce the scenario to just two countries, Japan & South Korea. The best payoff, again, is when both countries cooperate (with the ban), no whales right now perhaps, but plenty of whales forever, enough bounty for all, never an extinction. However if Japan chooses to cooperate ... and South Korea 'defects' and continues to whale, Japan gets nothing... nothing now, probably nothing later when the whales are extinct. If, on the other hand, Japan chooses to 'defect' ... then, even if South Korea defects, Japan will at least get something, not as much as if both had cooperated, and slimmer pickings now because they are BOTH competing for what fewer whales are available ... but still Japan will get something. And if South Korea cooperates while Japan 'defects,' that's even more whales for Japan to hunt.

Said another way, if Japan cooperates, it's either the best outcome ...or nothing. But if Japan defects, there's a pretty good outcome either way. Given Japan cannot control what South Korea will choose to do, Japan's best strategy is to defect. Given South Korea cannot control what Japan will choose to do, South Korea's best strategy is also to defect. Both will (eventually) choose to defect. Especially if one sees the other continuing to defect, for then it's known that the best outcome will never have the opportunity to occur. Even though, again, the best payoff for both would be if they cooperated and respected the ban.

Game theory explains the above line of thinking in more mathematically terms, assigning appropriate relative numerical payoffs for both countries under the four possibilities that result from each choosing to cooperate or defect. But it's easy enough to understand the concept intuitively as it is described.

Prisoner's Dilemma is the same sort of problem, in slightly different guise:

Two individuals are in prison without bail, suspected of attempting a major burglary. However, there is not enough evidence to convict them of that attempt, unless one of the two confesses, rats. There IS enough evidence to convict them of the lesser crime of trespassing. The two prisoners have been kept separate from one other, incommunicado. Detectives offer to each of them a deal in an attempt to get at least one of them to flip/snitch/rat ... if one prisoner 'rats' that they were both attempting a burglary, and the other does not rat, he will be given immunity for the crime and his confession will convict the other, who will bear the entire burden of the sentence for the more serious crime. If both 'rat,' there is no immunity for either, and they will share the burden of the sentence for the more serious crime. Of course, if neither 'rats,' they will both only be convicted of the trespassing that bears only minimal sentencing.

The scenario is structured in the same manner as the 'tragedy of the whales.' The best total outcome for both prisoners is to not 'rat' so that they will both get minimal sentences for trespassing. It is in both of their interests to 'cooperate with one another' and not 'rat/defect' the other out. However

¹ Game theory was invented in the 1950s by Professor John Nash, portrayed in the biopic "A Beautiful Mind." Cooperative non-whaling and mutual whaling defection are referred to as "Nash equilibriums," where both players will not change their strategies so long as the other player does not change. There is no such equilibrium where one cooperates and the other defects; sooner or later the cooperator will be drawn to defect as well.

choosing to cooperate is the 'all or nothing' option: either the trivial sentence for trespassing or, if the other 'rats,' the full burden of the sentence for the attempted burglary. But with rratting, there's a pretty good outcome either way ... either total immunity or the shared burden of sentencing for the attempted burglary. So the logical prisoner will rat.

Many societal groups are aware of these scenarios where the logical choice of individual defection defeat optimal group outcomes ... even though they understand them more so from a pragmatic perspective rather than a logical, game-theoretic perspective ... and so they have evolved rules, obligations and/or social norms that deter individual defection toward the greater good for the group.

For many decades in the early and mid 20th Century, the strict code of silence known as *omertà* prevented anyone in organized crime from defecting to law enforcement. *Omertà* is enforced by fear (of death), the stigma of being ostracized, and the loss of peer respect, the latter upon which members placed extraordinarily high value. In fact *omertà* is considered to be a 'code of honor.'

In the same vein, police officers often practice what is termed the "blue wall of silence" or "the blue shield," refusing to defect against other officers when they act unethically or illegally. Perhaps they do not fear death, but they are certainly avoiding being ostracized and losing peer respect. Also, an officer recognizes that his/her job is so difficult that at some point, he/she will succumb to an error in judgment. To enjoy the protection of other officers not defecting against him/her when that happens ... he/she must also refuse to defect regarding the errors in judgment of other officers. Medical doctors also informally adopt the same practice, refusing to ever criticize or defect upon other medical doctors, for the very same reason. Orthodox Jews, technically, are forbidden to defect to outsiders regarding inappropriate actions committed by other Orthodox Jews; one who does has committed 'mesirah' (the action of informing) and could be subject to death. The same phenomenon has been reported in that professional athletes will not defect regarding their awareness of the common drug 'doping' of other athletes.

Research has detected manifestation of the same phenomenon. Co-workers as subjects in a PD scenario have been found to resist defection against their own co-workers ... while they will more readily defect against unknown players. The existing relationships between them, as well as their familiarity with one another, and their need to be able to work together collegially in the future, 'bonds' them to an extent that diminishes their willingness to defect. Such social cohesion tends to increase cooperation. In fact, some German researchers eventually thought to study PD outcomes with female prisoners as subjects. Despite the best individual strategy of defection, the female prisoners chose cooperation about 57% of the time. They tended to work together, given the bond of their shared plight of imprisonment.

But when there is no type of 'bond' between the players, then it's 'every man for himself,' and they will defect away from what could have been a much better outcome had they cooperated together.

Zeno's Paradoxes

(authored January 25, 2020)

These are also of ancient origin, attributed to the Greek philosopher Zeno of Elea, during the 4th Century BC. There are three of them, but they are all logically equivalent, so only the first is discussed.

A man starts walking toward his destination, a mile away. Eventually he reaches the half-way point: one-half-mile walked, one-half-mile still to be walked. From that point, eventually he reaches the half-way point of the *remaining* distance: three-fourth of a mile walked, one-fourth of a mile still to be walked. Eventually he reaches the half-way point of *that remaining* distance: seventh-eighths of a mile walked, one-eighth of a mile still to be walked. And so it goes. The man can never, in all of eternity, reach the end of his one mile walk ... because he will forever be only reaching the half-way point of his remaining distance.

Mathematicians will claim that the eventual invention of calculus resolved the matter, in that it proves that the sum of the infinite series $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$ equals 1. That being true, still, the average person won't view that statement as much of a solution. "How could you ever add up a bunch of numbers that go on forever? And even if you could, how does that fact ever help the man get past his half-way point?" True enough.

The situation seems paradoxical because of the way it is being presented and considered; when doing so in a 'backwards' manner, thinking about the entire mile rather than all its pieces, the matter becomes much more clear and much less paradoxical.

A man can certainly walk a full mile, it happens all the time. And as part of walking that mile, he will first walk half that mile, then a quarter of that mile, then an eighth of that mile, and so on. So it's easy to see, without any use of calculus, that the full mile adds up to the infinite sequence of halving.

Of course, again, that doesn't explain how the man ever gets past these half-way points. Again, think about the entirety. Suppose the mile was one long piece of wood and we send out a carpenter to cut it in half. Then we ask the carpenter to cut one of the halves in half, resulting in two one-quarter pieces. Then we ask the carpenter to cut one of the quarter-pieces in half, resulting in two one-eighth pieces. Then we ask the carpenter to do the same thing, over and over again, until he is done. When will he be done? Never, of course. Because it would take him an *infinite amount of time* to do so; he would have to cut forever in order to finish.

The carpenter could do it, if only there were enough time. And in the same vein, the runner could do it, if only there were enough time. In the case of the runner, there IS enough time. Because to reach each half-way point, he will require only half as much time. If he is running at 's' miles per hour, then covering half a mile will require only $\frac{1}{2}s$. The total amount of time required to run the mile then is the sum of the infinite series $\frac{1}{2}s + \frac{1}{4}s + \frac{1}{8}s + \dots$ which, as we already know from above, totals to 1s.

While the runner appears to face an infinite series of half-way points (and half-way times) that are insurmountable, our now-better understanding of how infinite series behave reveal that he can indeed transverse an infinite series of both distances and times ... completely and in a finite amount of time.

It is our unintuitive perception that a infinity summation cannot converge into a finite total that here creates the illusion of paradox. Said another way, yes of course we can keep dividing a mile in half forever ... but it is still a finite mile that can be transversed in a finite time.

Hilbert's Hotel

(authored February 2, 2020)

Another supposed paradox related to the concept of infinity was first presented by German mathematician David Hilbert in 1924. I personally first encountered it around 1981 as it was presented in *White Light*, a science fiction novel written by mathematician Rudy Rucker.

A traveler arrives at Hilbert's Hotel. He does not have a room reservation. And he did not think he would need one, since the hotel has an infinite number of rooms. The front desk clerk informs him that, unfortunately, there are no vacancies, since the hotel is presently accommodating an infinite number of guests, one per room. The traveler proposes a solution. Move the guest in the first room to the second room, move the guest in the second room to the third room, move the guest in the third room to the fourth room. And so on. There will be enough rooms for all of the existing guests ... since there are an infinite number of rooms. Now the first room is left vacant for the traveler to occupy.

The traveler returns the following week with an **infinite** number of guests, none of whom hold a reservation. But, again, there are no vacancies, since the hotel is accommodating an infinite number of guests, one per room. The traveler insists they can all be accommodated. Move the guest in the first room to the second room, move the guest in the second room to the fourth room, move the guest in the third room to the sixth room. And so on. There will be enough rooms for the existing guests (using all the even-numbered rooms), leaving the infinity of all the odd-numbered rooms vacant for the traveler's entire infinite party to occupy.

It all seems impossible, right? An infinite number of guests matched to an infinite number of rooms, well that means there's no empty room. Yet the traveler has explained how to 'rearrange things' so that there is still one vacant room, or even an infinite number of vacant rooms.

The behavior of a finite set of numbers differs from the behavior of an infinite set of numbers ... and the behavior of the latter is not intuitive. That includes the size (or 'count-ability') of them.

Indeed the number of even-numbered rooms in a finite-sized hotel is half the number of total rooms in the hotel.

But this is not true regarding the number of even-numbered rooms in an infinitely-sized hotel. In that case, the count of the even-numbered rooms is actually EQUAL to the count of the total number of rooms. The two infinite sets share the same degree of ability to be counted. Said another way, the infinite set of integers can be mapped one-for-one to the infinite set of even-numbered integers (that being how we can 'count them.'). What's not working out here is that mapping the set of integers, in order to count and/or deduce the size of a set of numbers, doesn't work when mapping onto an infinite set.

There's more. Not ALL infinite sets share the same degree of 'count-ability.' The infinite set of 'real numbers' ... the set of rational numbers (numbers that can be written as the division of two integers without remainder) and the set of irrational numbers (numbers that cannot be written so, such as π , with its division never-ending) ... cannot be counted by integers. Said another way, the set of integers cannot be fully mapped onto the set of rational numbers. If we try to do so, mapping one-for-one, we can examine how the mapping 'counts' and logically deduce a bunch of rational numbers that were not 'covered' by the mapping.

To keep the example simple, consider the particular infinite set of rational numbers that fall between zero and one, inclusive. Suppose our counting, our mapping of the integers, started out with a random-sort of mapping like this:

#1: .498267...
#2: .584356...
#3: .872456...

And we just mapped forever, until every integer is mapped, one-for-one, onto every rational number we can find. We can *still* locate new rational numbers that are *not* mapped, not in the list.

#1: .498267...
#2: .584356...
#3: .872456...

That would be rational numbers that do NOT have a 4 as its first digit, do not have an 8 as its second digit, do not have a 2 as its third digit, and so on. Those numbers (and there's plenty of them) couldn't possibly be in the list that was mapped to the integers.

That means that the infinite set of rational numbers seems to be much LARGER than the infinite set of integers. And, in a way, that should sort of sound right: the infinite set of integers is *just* integers, but the infinite set of all rational numbers (say, from zero to infinity) 'feels' as if it contains so much more.

Said another way, there appears to be different 'sizes' of infinity! Amazing!

When dealing with a finite set of numbers, such as the number of rooms in an *actual* hotel, we can use the mapping of integers onto them in order to count them, in order to determine its size.

However mapping integers to an infinite set does not lead to the 'counting' and 'sizing' we would hope to achieve. Instead it leads to conclusions about size that appear bizarre.

That is because, by its very nature, infinity cannot be counted, cannot be sized. The entire notion of size only applies to finite quantities. And that concept, when thought about in that way, is really just common sense.

#