The Solution to the Raven Paradox

Ruadhan O'Flanagan

October 2009

Abstract

The raven paradox goes as follows: "All ravens are black" is equivalent to "All non-black things are non-ravens". Since the observation of a non-black non-raven provides evidence or support in favour of the latter, that same observation must provide the same support for the former. This leads to the conclusion that the observation of a white shoe, for example, provides evidence or support for the proposition that all ravens are black. The solution to the paradox is given here.

Asymmetry

$x = y$ and $y = x$ are not equivalent in computer programming. The first sets the value of x; the second sets the value of y. The solution to the paradox rests on this asymmetry.

Predication

For a person who does not know John's height but does know Mary's height, the sentence, "John's height is the same as Mary's height" serves as an instruction to mentally set the value of the variable "John's height" equal to the value of the variable "Mary's height." This is comparable to executing a command such as $x = y$.

Even in the case when neither Mary's nor John's height is known, the command to set the (estimated) value of "John's height" equal to the (estimated) value of "Mary's height" has a different effect to the command to set the value of "Mary's height" to the value of "John's height". For example, we might know that John is a man and that the average man's height is 1.8 metres while we also know that Mary is a woman and the average woman's height is 1.7 metres. If we set Mary's estimated height equal to John's, we will end up in a situation where we estimate both to be 1.8 metres tall. If we set John's estimated height equal to Mary's, our estimate will be 1.7 metres for them both.

The asymmetry is present in the English sentence, "John is as tall as Mary". "John" is modified by the sentence; "Mary" is not. Mary is mentioned only for reference, so that it can be known what is attributed to John. In the terminology of computer programming, the variable "John's height" was assigned the value that was stored in the variable "Mary's height". In the terminology of grammar, "as tall as Mary" was predicated of "John".

To be sure, there is a sense in which "John is as tall as Mary" is equivalent to "Mary is as tall as John", namely as conditions. One condition is true if and only if the other is true. However the sentences are not merely conditions. They convey additional information, namely which part is to be modified and which part is for reference so that it can be known what modification to make.

"The raven is black" is likewise naturally interpreted as an assignment of the value "black" to the variable "the raven's colour". To interpret it as an instruction to modify the predicates of "black" and leave the predicates of "the raven" unchanged would be highly unusual and contrived. It is not even plausible that the predicates of "black" and those of "the raven" are intended to be modified in equal amounts.

Contrapositive

An if-then statement in a computer programming language has both a condition and a command. The if-then statement is a command to check the condition and execute the command if the condition is satisfied.

In the sentence, "If the object is a raven then the object is black", the part which says "the object is a raven" is a condition, while the part that says "the object is black" is an assignment of the value "black" to the variable which represents the object's colour. Alternatively it could be viewed as an assignment of the predicate "true" to the proposition "the object is black".

Naïve symbolic logic freely substitutes "B or not-A" in place of "If A then B". In doing so, it loses that part of "If A then B" which is captured by if-then statements in computer programming. Naïve logic uses *propositions*, which are objects that can act as either conditions or assertions, which are a type of command. In natural language and computer programming, the B in "If A then B" can be a command, such as "If the resource has run out, then get some more."

Boolean variables can take the values TRUE and FALSE. These are convenient because, like propositions, they allow us to deal with objects which can serve as both conditions and commands. Introducing r and b to represent the propositions "the object is a raven" and "the object is black" respectively, we can write the "contrapositive commands":

- IF $r = \text{TRUE}$ THEN $b = \text{TRUE}$
- IF $b = \text{FALSE}$ THEN $r = \text{FALSE}$

The first sets b to TRUE if the value of r is TRUE. The second sets r to FALSE if b is FALSE.

That these have different effects can be seen from the fact that, if the initial condition is $r = \text{TRUE}$, $b = \text{FALSE}$, then executing the first command will result in $r = \text{TRUE}$, $b = \text{TRUE}$ while executing the second command will result in $r = \text{FALSE}$, $b = \text{FALSE}$.

This proves that, although contrapositives are equivalent in naïve symbolic logic, contrapositives are not equivalent when interpreted as conditional assignments.

Paradox Resolved

When one acquires new knowledge or begins to believe something that one previously did not, there is a change of inner condition which occurs, comparable to the change of state which occurs when a computer executes a command. If one previously did not know that the object X had the property Y, and then subsequently learns that X does indeed have the property Y, then the status of X with regard to the property Y changes from "undetermined" to "true".[1]

This occurs also with induction. When one has observed that many objects with the property X also have the property Y and no counterexamples have been encountered, one might begin to believe that *all* objects with the property X also have the property Y. Setting aside the question of whether this is justified or advantageous or even whether people actually do it at all in the course of reasoning, we can examine the structure of the inductive act and ask what changes it makes.

The inductive act is clearly a predication; there is a class of objects, those with the predicate X, and Y is predicated of the objects in that class. "All

[1]Here, "property" is used to refer to characteristics of the object itself, while "predicate" is used to refer to the grammatical or logical symbol which a speaker or reasoner might attribute to an object to represent that the object has that property.

X are Y." This produces a *change* in the status of every hypothetical object, O, effectively executing the command

- IF "O is X" = TRUE THEN "O is Y" = TRUE

That is the rôle of the expression "All ravens are black" in the inductive act. It is *not* a condition, namely something which can be true or false and which involves no change. It is an *action*, executed like a command, and it makes a change to the predicates of objects which are ravens, but does not change which objects are designated as ravens and which are not.

For example, a person who has observed one thousand ravens and found them all to be black might, before going through with the induction, say something like, "I estimate that there are a millon ravens in the world, and I know that at least one thousand of these are black, while I do not know what colour the remaining ravens are, and I have not yet speculated about that." After the induction, he might say, "I now believe that the estimated 999,000 ravens that I have not yet examined are black, having spent time speculating about their colour while referring to the collected data." Whether it is a wise thing to do or not, that is what occurs when one generalizes from "The ravens I have seen are black" to "All ravens are black".

To see that the inductive action which predicates "black" of "all ravens" is not the same as the inductive action which predicates "non-raven" of "all non-black things", consider abstract predicates X and Y and suppose that nothing at all is known about X or Y or their relationship to each other or how many objects might have those predicates. In such a condition of ignorance, there is no information which might favour X over not-X, so the probablity that a randomly chosen object will have the property X is one half, and the same applies for Y. Considering a population of objects, one would estimate that one quarter of the objects would be both X and Y, one quarter would be X and not-Y and so on.

Now consider how the estimates would change if the predication "All X are Y" were to be executed. The half of the population which were estimated to have the property X would then all be assigned the predicate Y, while nothing at all would be predicated of the objects which were not-X. The final estimate of the fraction of objects with property Y would be three quarters, which includes all of the objects estimated to have property X and half of those estimated to have property not-X.

The effect of the predication, "All X are Y", then, is to increase the estimate of the fraction of objects with the property Y, while leaving the estimate of the fraction of objects with property X unchanged. "All non-Y are non-X" would have a different effect, namely to predicate not-X of certain things, thereby increasing the estimate of how many things are non-X.

The inductive action which predicates “black” of “all ravens” therefore has a different effect from the action which predicates “non-raven” of “all non-black things”, since the former changes the estimated fraction of the objects which are black and the latter changes the estimated fraction of objects which are non-ravens.

Since the two inductive actions are not equivalent, evidence or support for one is not generally evidence or support for the other.

This might be compared to computer programming by saying that setting the value of y equal to the value of x by executing the command $y = x$ is not equivalent to setting the value of x equal to the value of y by executing $x = y$, and a reason or motivation or justification for executing one command is not necessarily a reason or motivation or justification for executing the other command.